Condition-based Maintenance:

Capital goods industry

Hao Peng

*Eindhoven University of Technology, the Netherlands*

**Abstract**

*Condition-based maintenance (CBM) is one type of preventive maintenance policy. CBM has attracted lots of attentions of both academia and industry due to the development of advanced sensor technology and measurement devices. The proper implementation of CBM can reduce the frequency of random failures and the expected cost of maintenance during the lifecycle of a system. In this chapter, a brief overview of different maintenance strategies is first provided for the readers who are not familiar with maintenance optimization models. Then several elementary models about CBM will be introduced to help the readers get a general idea of the optimization models in this field.*

Keywords: reliability theory, failure time distribution, corrective maintenance, age based maintenance, block replacement, minimal repair, delay time model, random coefficient model

INTRODUCTION

In capital goods industry, the availability of systems is crucial to provide the needed functions or services to customers/users. Aircraft, trains, wafer-steppers, and magnetic resonance imaging (MRI) scanners are examples of such systems. The inconvenience of trains and/or aircraft not running when needed (and planned) is a great inconvenience to travellers, but also a significant loss of revenue for airlines and railway operators. Wafer-steppers are used in the bottleneck production step of semi-conductor manufacturing. When a wafer-stepper is down, it causes the standstill of an entire semi-conductor factory. For the case of wafer steppers, these costs are in the order of magnitude of 100000 EURO per hour. The unavailability of MRI-scanners is perhaps the most costly as it can lead to the loss of human life under some circumstances. All these examples illustrate that keeping capital assets up and running is of critical importance.

On the other hand, the cost of maintenance to keep the systems up and running is very high. Oner et al. (2010) estimate that the costs of maintenance and unavailability of a capital asset over its lifetime (typically one to several decades) is typically three to four times the acquisition cost of capital assets. In fact, in 2003 spare part sales and services (mostly maintenance) accounted for 8% of the gross domestic product in the United States (AberdeenGroup, 2003). More recently, US bancorp estimated that the yearly expenditure in the US on spare parts amounts to 700 billion dollars which is 8% of the US gross domestic product (Jasper, 2006).

How to create optimal maintenance policies and schedule the tasks of maintenance efficiently becomes a challenging problem, since we not only need to reduce the total cost of ownership but also have to achieve a high level of system availability and reliability. Many maintenance strategies have been proposed to tackle this problem. Figure 1 gives an overview of maintenance strategies. Different from regular production operations, maintenance operations are not instigated by demand from an outside customer, but by the need for maintenance of equipment. The needs of maintenance can come from machine degradations, failures, or technology upgrades. If a maintenance activity is initiated by technology upgrade, it has been defined as modificative maintenance. Modificative maintenance concerns interchanging a part with a technically more advanced part in order to upgrade the equipment. This form of maintenance is usually project based and non-recurring.

The maintenance strategies that occur most often are preventive and corrective maintenance. Under a corrective maintenance strategy, a part is not replaced until it has failed, while under a preventive maintenance strategy, it is possible to replace parts before unexpected failures occur. To determine which strategy should be used, the information about reliability behaviors of systems plays an important role. Corrective maintenance is an attractive option for parts that do not wear, such as electronics. For parts that do wear, it can be beneficial to follow a preventive maintenance strategy.

Preventive maintenance strategies can be further divided into time, usage and condition based maintenance. Under time based maintenance, the ages of systems or calendar times will be the variables that determine the executions of maintenance activities. For example, an age based maintenance policy will optimize the age limit after which a system will get preventively maintained. Block replacement policy has a fixed maintenance interval to optimize, which specifies the time points at which a system will get preventively maintained. Notice that the information about reliability behaviors, which is used to optimize a certain time based maintenance policy, is the failure time distributions of systems/components estimated from historical data or expertise knowledge.

For some cases, time in the field is not the most appropriate measurement of system degradation toward failure, e.g., airplanes (Tinga, 2010). Usage/load based maintenance has been proposed to assess the system condition (and prognostics), by applying usage or load parameters in a physical model that describes the physics of failure. The usage or load can be recorded or monitored during service. For example, the start-stops and rotational speeds can be recorded for a plane; the temperature, strain and electrical current can be monitored for mechanical and electrical systems. Based on these records of usage/load over time, the system condition will be evaluated and preventive maintenance actions will be performed when the system condition is close to a failure status. The predictions of failures under this maintenance strategy can be much more accurate than those of the time based approaches, since a physical model is used to calculate the system degradation.

In condition based maintenance, the actual condition of a part is gauged and maintenance is conducted based on this. The condition of a part can be measured either by inspections (periodically or according to a certain schedule) or through continuous monitoring of sensors. How the condition of equipment is measured depends on the nature of equipment, measurement accuracy or economic concerns. The condition of ball bearings can be measured by monitoring the amplitude of vibrations around the bearing through sensors. The condition of a metal part in a structure can be determined by visually inspecting the number and length of cracks periodically. For mechanical systems with moving parts, the concentration of ferrous parts in the lubrication fluid can be inspected periodically as an indication of the wear.

The major difference of CBM compared with other maintenance strategies is the utilization of the advanced information about the health status of a component or a system (collected through inspections or sensors). For condition-based policies, systems or components have multiple intermediate states in between the failure state and the perfect/newest state. The transition of degradation states can be described by many different types of probability models, e.g., delay time model, or Markov process. Or, the deterioration of systems may follow a continuous stochastic process, e.g., random coefficient model, gamma process, or Wiener process. The advanced information collected through inspections or sensors will be used to estimate the parameters of a selected probability model for a degradation process. Based on the probability models of degradation processes, the inspection and replacement/maintenance decisions can be made to optimize the CBM policies from the cost and reliability/availability perspectives.

Maintenance operations are subject to considerable uncertainty due to the randomness of systems’ reliability behaviors. There is uncertainty both with respect to timing (When will maintenance /replacement be needed?) and content (What parts need maintenance /replacement?). Corrective maintenance cannot be planned for ahead of time. Therefore, both the timing and content of maintenance operations are uncertain, which gives difficulties for the scheduling of maintenance resources, i.e., human resources, repair tools or spare parts. For preventive maintenance, since the aim is to repair or maintain a system before the unexpected failures, a certain portion of the maintenance activities is planned in advance. For instance, under time based maintenance strategy, the preventive maintenance actions will be scheduled based on the ages of systems or the calendar time. Therefore the timing and content uncertainties of preventive maintenance strategies are reduced to a certain extent compared with corrective maintenance. The maintenance resources for preventive maintenance can be utilized more fully than for corrective maintenance.

Under usage based maintenance strategy, the preventive maintenance actions will be scheduled based on the usage profiles of systems. Similarly, for condition based maintenance with sensor monitoring, the preventive maintenance actions will be scheduled based on the system conditions. Since the degradation usually is a stochastic process, the timing of preventive maintenance actions for a system under usage based or sensor based maintenance strategy is unknown and is dependent on the usage or the condition of the system. This will make the timing of preventive maintenance actions uncertain, but the content of replacements is normally known due to the sensor monitoring. The parts that need replacement can be inferred from the sensor signal. Notice that for sensor based CBM, remaining useful lifetime (RUL) is usually estimated from the sensor data, by using all kinds of statistical approaches (Si, Wang, Hu, & Zhou, 2011). If this is the case, age based maintenance models can be used to determine the optimal replacement time point, which eliminate the uncertainty of maintenance timing.

For inspection based CBM, inspection schedules are specified in advance. Therefore, the inspection time points become the opportunities or moments to perform preventive maintenance actions. Then the timing of maintenance activities is known for the preventive maintenance actions. However, the content of maintenance depends on the inspection outcomes. If a system or component is found to be defective or in an unacceptable degradation state, preventive maintenance actions will be performed immediately.

The proper implementation of CBM can reduce the frequency of random failures and the expected cost of maintenance during the lifecycle of a system. To obtain these benefits from CBM, we need not only accurate prediction models, but also optimization models to smartly determine the inspection and replacement policies for single-component systems, as well as to smartly coordinate the maintenance activities for multi-component systems. In this chapter, the major objective is to provide the conceptual and theoretical foundations for the use of CBM elementary models in capital goods industry. This book chapter is to fill the need for an introductory text on CBM modeling for scientists. We intend it to be useful in the engineering field; any field in which systems of degrading and observable entities are of interest.

This chapter has these objectives:

1. Introduce several elementary models of maintenance optimization for single-component systems.
2. Introduce several elementary models of CBM for single-component systems.

Background

CBM has attracted lots of attentions of both academia and industry due to the development of advanced sensor technology and measurement devices. In a CBM framework, after several key steps, i.e., data acquisition, data processing, diagnostics and prognostic, maintenance policies will be optimized to minimize the operational costs or maximize the availability of systems (Jardine, Lin, & Banjevic, 2006). The main difference between the conventional maintenance models and CBM models is the utilization of condition measurements. Failures usually occur when the degradation level of a system reaches its failure threshold level, so that the condition monitoring data and stochastic models of the degradation processes are often necessary to estimate remaining useful lifetimes (RUL) or reliability functions. Si et al. (2011) distinguished two types of probability models of RUL estimation: directly observed CBM models (e.g., regression-based models, Wiener process, Gamma processes, Markovian-based models) and indirectly observed CBM models (e.g., stochastic filtering-based models, covariation-based hazard model, hidden Markov model).

For single-component systems under continuous monitoring in the industry of advance capital goods, if we can predict the RUL, the optimization of CBM policies can be formulated similarly to the age based maintenance policies. Gebraeel et al. (2005) extended the general degradation model to estimate the RUL distribution from sensor signals, using a Wiener process and Bayesian updating. Using this technique, a single-unit replacement problem is formulated as a Markov decision process to develop a structured replacement policy (Elwany, Gebraeel, & Maillart, 2011).

If inspections are needed to detect the system state, based on the general random coefficient model (Lu & Meeker, 1993), Wang (2000) proposed a CBM model to determine the optimal control limit and the inspection interval in terms of costs, downtime or reliability. For monotonic stochastic deteriorations, a Gamma process can be used for condition-based maintenance optimization (Van Noortwijk, 2009). The CBM models in this case were developed to have a single-level control limit or a multi-level control limit, under the scenarios of periodic inspection, aperiodic inspection or continuous monitoring. Renewal theory can be used to evaluate the expected total cost rate.

If the degradation process could be modeled as discrete states, Markovian-based models were applied. The optimal replacement policies were derived from observable Markov processes (Kharoufeh, Solo, & Ulukus, 2010) or the evolution of the hidden states (Jiang, Kim, & Makis, 2013). Markov decision process is an analytical tool to formulate the problem. Moreover, Proportional Hazards Models are also often used to relate the system's condition variables to the hazard function of a system (Vlok, Coetzee, Banjevic, Jardine, & Makis, 2002), so that the maintenance policies can be optimized with respect to the optimal risk value of the hazard function.

In this chapter, we mainly demonstrate the evaluation of average total cost rate by renewal theory for single-component systems. The degradations of the systems are assumed to be directly observable through periodic inspections.

Basics of reliability theory

Since CBM is a branch of research in the field of reliability theory, we will first review some basic results from reliability theory. Not all of the results from reliability theory are relevant to CBM. Thus in this chapter we only present the results that will be used in the following sections. To get a more comprehensive introduction to reliability theory, the readers can refer to the book of Ebeling (2009).

Failure Time Distribution

Present your perspective on the issues, controversies, problems, etc., as they relate to theme and arguments supporting your position. Compare and contrast with what has been, or is currently being done as it relates to the chapter's specific topic and the main theme of the book.

Failure time distribution provides the basic information that is needed to establish optimization models for time based maintenance strategies. Because some of the CBM models are closely related to the time based models, we will provide the basic reliability theory about failure time distribution in this subsection.

In reliability theory, we mainly deal with non-negative random variable. Let random variable *T* denotes the time until failure of some component. If *T* is a continuous random variable, we denote its distribution by FT(t) = P(T ≤ t) and assume it has a density f T (t) = d FT (t)/dt. The reliability of the component at time t is the probability that component survives beyond time t and is denoted by R(t).



The mean time to failure (MTTF) is the expectation of T,



The second equality in the above equation holds because fT(t)=d FT(t)/dt=-d RT(t)/dt, and the final equality holds because when t approaches to infinity R(t) approaches to zero.

The failure time distribution describes the likelihood of failure on every possible time point over the entire population of a component or system. In some cases, we also need to know the conditional failure time distribution given the fact that the component has survived for a certain period of time. The condition failure time distribution can be derived from the failure time distribution,



We may also want to know the conditional probability of failure in a certain interval ε given that the component has survived for a time period t,



(1)

Based on this conditional probability of failure, we could develop a concept called failure rate or hazard rate (we will refer to it as the failure rate), which is similar to the concept of failure time density fT(t)=d FT(t)/dt. In developing the failure time density, we look at the probability of failure in a certain interval (t, t+ε), whereas in developing the failure rate we look at the conditional probability of failure in a certain interval (t, t+ε). Therefore, the failure rate h(t) is given as



The second equality is by substituting equation (1) into h(t), and the third equality holds because of the definition of fT(t). This failure rate h(t) can be loosely interpreted as the instantaneous expected number of failures per time unit at time t. The failure rate reveals some essential features about the degradation process of components. If the failure rate is an increasing function of time, the component degrades over time. The component is more and more likely to fail when it ages. Mechanical devices typically have an increasing failure rate. For instance, when the cracks of structure materials grow, the structure materials will be more vulnerable to failures. If the failure rate is a decreasing function of time, the component becomes more reliable over time. This often happens when the component or system is at an early stage of operation or in a burn-in period. For example, a transmission system in a car needs a time period to burn-in and let the components run smoothly with each other. Another well-studied type of failure rate function is constant failure rate. It means the component has the same likelihood of failure over time, which implies the component has no memory on its age or its age does not have any effect on the component. It often reflects the component is subject to purely random failure events. Electronic components often have constant failure rate, since electronics are usually not subject to wear. Then the failures are often due to unpredictable events, e.g., random shocks from the environments.

In most of the cases, the failure rate is not a monotonic function over time. It has been discovered through empirical studies that the failure rate function has a bathtub shape, as shown in Figure 2. At the early stage, due to the infant mortality features, the component will experience a decreasing failure rate; in the middle of the life cycle, since the burn-in period has been passed and the wear-out hasn’t started yet, the component is subject to purely random failure events and the failure rate is constant; at the end of the life cycle, because of the degraded status, the reliability behavior of the component will get worse over time.

All of the functions discussed before, i.e., distribution function FT(t), density function fT(t), reliability function R(t) and the failure rate function h(t), are derived based on the failure time distribution. Therefore, based on any one of the four functions we can obtain the other three functions. In Table 1, the one-to-one relationships among the four functions are given to show that how to obtain a certain function from another function.

Table 1. One-to-one relationships between various functions about failure time distribution

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | fT(t) | FT(t) | R(t) | h(t) |
| fT(t) |  |  |  |  |
| FT(t) |  |  | 1-FT(t) |  |
| R(t) | - | 1-R(t) |  | - |
| h(t) |  |  |  |  |

The failure time distribution can take many different forms. For instance, exponential distribution, uniform distribution, Erlang distribution, Gamma distribution, lognormal distribution and Weibull distribution, are the widely used failure time distributions. For detailed introductions of these distribution functions the readers can refer to the textbook written by Ebeling(2009).

Corrective Maintenance Policy

The corrective maintenance policy is to replace/maintain a component every time it fails. Under this maintenance policy the preventive maintenance actions will not be taken or planned. The average cost rate of maintenance for this policy can be evaluated by renewal theory if we assume the replaced component is new and from the same population as the failed component, and the replacement time is negligible compared with the operating time. For the readers who are not familiar with the renewal theory, a technical note on renewal theory is provided below in the subsections. The costs up to time t for the corrective maintenance policy are a renewal reward process. The expected renewal cycle length for this corrective maintenance policy is the mean time to failure, E[T], since we will not preventively replace the component before the unexpected failures. The renewal cycle cost is a constant variable Cc because there’s no probability of taking other preventive maintenance actions. The average cost rate of corrective maintenance policy can be found by the renewal reward theorem, i.e., Cc/E[T].

An example is given to show the implementation of corrective maintenance policy. Suppose that the lifetime of a component is uniformly distributed from 10 to 15 time units and that the corrective maintenance action costs 100 euros. If we choose to apply a corrective maintenance policy for this component, the average cost per unit time is 100/((10+15)/2)=5.7 euro per time unit.

Renewal Theory

A renewal process is a counting process in which the times between events (also called renewals) are independently and identically distributed (i.i.d.). Let X1, X2,…, Xi be a sequence of non-negative i.i.d. random variables with common distribution F(x). Xi is the time between the (i-1)th and i th renewal; for example Xi might represent the time to failure of a component. Then the total number of renewals that have occurred up until time t, i.e., N(t), is given as



Such a process is called a renewal process and denoted by N(t). Figure 3 shows an example of a renewal process with the notation that we introduced.

Now suppose that there is reward (cost), Wi, associated with each renewal i. Furthermore, assume that W1, W2,…, WN(t), is a sequence of i.i.d. random variables with an expected value less than infinity. The total reward (cost) up until time t is denoted by Y(t) and satisfies



Y(t) is called a renewal reward process. Notice that we do not assume that Wi and Xi are independent. For such a renewal reward process, the average cost per time unit can be evaluated based on a theorem. For the readers who are interested in the proof of this theorem, they can be referred to Ross (1996) for a formal proof. The theorem tells us that the average cost per time unit of a renewal process satisfies

,

where E[Wi] and E[Xi] are the expected reward per renewal cycle and the expected renewal cycle length respectively.

Age Based Maintenance Policy

Under the age replacement policy, a component is replaced whenever it has been used for a fixed amount of time *A* or upon an unexpected failure before age *A*. The average cost rate of maintenance for this policy can also be evaluated by renewal theory if we assume the replaced component is new and from the same population as the original component, and the replacement time is negligible compared with the operating time. There are two types of renewal events for this age based maintenance policy at the end of a renewal cycle. One is the preventive maintenance action since the age of the current component reaches a time limit *A*; the other one is the corrective maintenance action due to an unexpected failure before its age *A*. The probabilities of the occurrences of the renewal events can be determined based on the failure time distribution fT(t). The probability of having preventive maintenance action at the end of a renewal cycle is thus given as

.

The probability of having corrective maintenance action at the end of a renewal cycle is given as .

Therefore, the expected renewal cycle cost is

,

(2)

where Cp is the cost of preventive maintenance action, and Cc is the cost of corrective maintenance action. Under the event of preventive maintenance renewal, the renewal cycle length is the age limit *A.* Under the event of corrective maintenance renewal, the renewal cycle length is the failure time. Therefore, the expected renewal cycle length is

.

(3)

The average cost rate of the age based maintenance policy is then ECC/ECL. The optimal replacement time *A\** can be found by minimizing the average cost rate. When *A* approaches to infinity, this policy is equivalent to the corrective maintenance policy.

To give an example, let us reconsider the component with Cc=100 and a lifetime that is uniformly distributed between 10 and 15 time units. Suppose that planned preventive maintenance for the component costs 30 euro. For a given replacement time *A*, we have (based on Equation 2)



(4)

And the expected cycle length is (based on Equation 3)



(5)

Substituting Equation 4 and 5 into the calculation of the average cost rate, we obtain the following expression for the expected cost per time unit,



The optimal age limit A can be obtained by checking the first derivative of the average cost rate function,



This first derivative is always positive in the range of (10,15). The average cost rate function is increasing in the range of (10,15). Hence, the optimal age limit A is at 10 to minimize the average cost rate. If we set A equal to 10, the average cost rate is 3 euro per time unit. Compared with the corrective maintenance policy proposed in the previous section, we saved (5.7-3)/5.7=47.4%.

Block Replacement Policy

In the age based maintenance policy, from the scheduling perspective, the maintenance manager needs to keep track of the ages of the components, i.e., the scheduler needs to record the calendar time of the previous replacement for each component under the age based maintenance policy. Normally the preventive maintenance times of different components are different due to the randomness of failures. Another way of management is to plan the preventive maintenance actions at fixed time points on calendar, r, 2r, 3r,… and an unplanned corrective maintenance is done if the component fails in between these time points. This policy with a fixed internal r can be attractive since the coordination of the preventive maintenance actions of different components is relatively easy.

For such a block replacement policy, unlike the age based maintenance policy, the time between replacements is not independent and identically distributed. For instance, consider a component that has just been preventively replaced at time 0. The time until the next replacement is distributed as min(T,r). Suppose that a failure happens at time x, which is smaller than r. Then the time until the next replacement is not distributed as min(T,r) any more, but has a distribution as min(r-x,T), which is dependent on the previous replacement time and not identical to the previous replacement cycle. Fortunately, renewal reward theory also applies to the block replacement policy by defining a cycle as the time between planned preventive replacements,



The expected cycle cost includes one planned preventive replacement cost Cp at the end of the cycle and the expected cost of corrective maintenance actions during the cycle. The expected cost of corrective maintenance actions equals the corrective maintenance cost Cc multiplied with the expected number of failures during an interval of length r. Let us denote the expected number of failures during r as MT(r). Within an interval r, we can ignore the existence of the preventive maintenance actions. Then the failure process is still a renewal process with independent and identically distributed failure times. For such a renewal process, the expected number of renewals up until time r, E[N(r)], obeys the integral equation provided in the subsection, i.e.,



where FT(.) and fT(.) is the cdf and pdf of the failure time T. Therefore the expected cycle cost is



The average cost rate depends on r and is given by ECC/ECL. The optimal block replacement interval r can be found by minimizing the average cost rate.

To give an example, let us reconsider the component with Cc=100 and a lifetime that is uniformly distributed between 10 and 15 time units. Suppose that planned preventive maintenance for the component costs 30 euro. Since 10≤T≤15, the optimal r should also be in the range of [10,15]. For instance, if r is smaller than 10, there will be no random failures in between the preventive maintenance actions, but the lifetime of a unit is not fully used, which is not desirable. If r is larger than 15, there will be one random failure at least in between the preventive maintenance actions, which is also not desirable. To determine ECC, we need to find MT(r) for r ∈[10,15]. The number of failures in an interval with a length less than 15 can be at most 1. Therefore, for this example, MT(r)=FT(r) for t ∈[10,15]. Then the expected cycle cost is



The average cost per time unit is

.

The first derivative of this objective function is

.

Since the first derivative is larger than zero, the average cost rate is thus increasing in the range of [10,15]. Therefore the optimal block replacement interval r=10 and the minimum average cost rate is 3 euro per time unit, which is the same as the age based replacement policy.

Renewal Theory

For a renewal process, let N(t) denote the total number of renewals that have occurred up until time t. The expected number of renewals till time t is denoted as M(t). Then the renewal function M(t)=E[N(t)] satisfies



where F(.) and f(.) is the cdf and pdf of the random time between the renewal events.

Proof: The proof follows from conditioning on the random time of the first renewal event X1.



The third equality follows because N(t)=0 if X1>t. The fourth equality follows because if X1=x<t, at least one renewal occurs, and the expected number of renewals from x to t is M(t-x).

To computing M(t), we need to solve the above integral equation.

Block Replacement Policy with Minimal Repair

If we consider the cases for which failures in a block replacement interval can be fixed by minimal repairs, the evaluation of average cost rate will take a different form. Minimal repair means the machine will not be repaired to as-good-as-new state, but to a state that is statistically identical to the state just before failure. The related cost can be denoted as Cmr. In practice, minimal repairs are performed by using duct-tape, tie-wraps and other ad-hoc solutions to get a component functioning again, without actually replacing it or performing thorough maintenance. A perfect repair brings the failure rate back to time zero hT(0) after the replacement . By contrast, after minimal repair, the failure rate will be the same as the failure rate before the failure happens. Because of this, the expected number of failures in a block replacement interval is no longer given by MT(r). Recall that the failure rate can be loosely interpreted as the instantaneous expected number of failures per time unit at time t. Then the expected number of failures during a block replacement interval is

.

The expected cycle cost is

.

To give an example, let us reconsider the component with Cmr=20 and a lifetime that is uniformly distributed between 10 and 15 time units. Suppose that planned preventive maintenance for the component costs 30 euro. Since 10≤T≤15, the optimal r should also be in the range of [10,15]. For instance, if r is smaller than 10, there will be no random failures in between the preventive maintenance actions, but the lifetime of a unit is not fully used, which is not desirable. If r is larger than 15, there will be one random failure at least in between the preventive maintenance actions, which is also not desirable. To determine ECC, we need to find the expected number of failures in a block replacement interval for r ∈[10,15], which is



The expected cycle cost is



The average cost rate is



The first derivative is



In the range of [10,15], the first derivative is positive. Then the average cost rate is increasing. The optimal block replacement interval is 10 in this case. The average cost rate is 3, which is the same as the block replacement policy without minimal repair and the age based policy.

Delay time model

For age-based policy or block replacement policy, the information needed is the failure time distributions of components or systems. For example, in the calculation of age-based policy, we need to know the distribution function of the random lifetime of a component, FT(.). In the calculation of block replacement policy, we need to know the expected number of failures during an interval, MT(.), which is also derived based on the failure time distribution. By using the failure time distribution to describe the health status of a component or a system, we assume there are two states for a component or a system, i.e., the failure state and the working state. The random failure time T is the duration of the working state.

For condition-based policies, systems or components have multiple intermediate states in between the failure state and the perfect/newest state. The transition of degradation states can be described by many different types of probability models, e.g., delay time model (DTM), Markov process. Or, the deterioration of systems may follow a continuous stochastic process, e.g., random coefficient model, gamma process, Wiener process. Then the system state is continuous. Based on the information of degradation processes, the inspection and replacement decisions can be made to optimize the CBM policies. Renewal theory may be used to evaluate the expected total cost rate. Markov decision process is also an analytical tool to formulate the problem. In this chapter, we mainly demonstrate the use of delay time model in CBM.

The delay time model was first developed by Christer and Waller (1984). It assumes that a component or a system has three states: normal, defective and failed.

DTM1.png

As shown in Figure 4, the duration of the normal state, also referred to as the time to defect, is a continuous non-negative random variable denoted by X. The duration of the defective state, also referred to as the delay time, is a continuous non-negative random variable denoted by H. The cumulative distribution function and the probability density function of the time to defect are denoted by FX(.) and fX(.) respectively. The cumulative distribution function and the probability density function of the delay time are denoted by FH(.) and fH(.) respectively. Under the normal state, the system works fine and defects do not exist in the system. Under the defective state, the system still works but the defects of the system appear and can only be detected by inspections. The failure state of the system is self-announcing. The advanced information considered by DTM is the defective information obtained through inspections, compared with the age-based policy or block replacement policy. Given such information, we are interested in obtaining the optimal inspection and maintenance policy. The DTM is an abstract of many engineering systems. Of course, it is a simplified description of complex failure mechanisms. However for most cases, this simplification works pretty well for maintenance optimization.

For example, the condition of ball-bearings can be measured via the amplitude of vibrations around the bearing. After a certain period of operating, the condition of ball-bearings becomes less good and the amplitude of vibrations becomes larger. The engineers define a certain limit of vibration above which the engineers will see the system as ""defective". The system will stay in the defective state for a while till breakdown.

To give another example, the condition of a metal part can be determined by visually inspecting the number and length of cracks. After a certain period of operating, the condition of a metal part becomes worse. The engineers define a certain criteria for cracks to determine the system state as "defective". The system will stay in the defective state for a while till breakdown.

For metal systems with moving parts, the concentration of ferrous parts in the lubrication fluid is measured as an indication of the wear. After a certain period of operating, the concentration of ferrous parts becomes larger. The engineers define a certain limit for this concentration measurement above which the system is seen as "defective". The system will stay in this defective state for a while till breakdown.

A periodic inspection policy for a single-component system

We consider a periodic inspection policy for a single-component system that has a single failure mode (i.e., the component is subject to one dominant failure mechanism). We inspect the system with a fixed interval r. We assume the inspections are perfect, i.e., if the system is in the defective state the inspections can detect it immediately. Our maintenance policy is that if we detect the defective state upon inspections, we will immediately replace the component with a new one. If the component fails before the next scheduled inspection, we will perform a breakdown corrective maintenance action immediately. The maintenance time is negligible. Notice that there are two types of renewal events in this periodic inspection policy, i.e., the inspection renewal when defects are found and the breakdown renewal. After the system is renewed by a new component, the inspections are rescheduled from the renewal point. We assume the inspection cost is Ci. The preventive maintenance cost is Cp, which is the replacement cost. The corrective maintenance cost is Cc, which is the replacement cost plus the failure cost. We first consider a simple case that the time to defect X follows an exponential distribution. According to the memory less property of the exponential distribution, this implies an inspection renews the system regardless of whether a defect was identified or not. Since each failure or inspection renews the system with associated costs, the process is a renewal reward process, and the long term expected cost per unit time, CR(r), is given by,



where ECC is the expected renewal cycle cost and ECL is the expected renewal cycle length which is the interval between two consecutive renewals. Let us first look at the expected renewal cycle length. Starting with a renewal point, if the failure time X+H is larger than the inspection interval r, then the system gets renewed at time point r, i.e., the renewal cycle length is r. If the failure time X+H is smaller than the inspection interval r, then the system gets renewed at a failure, which is random. The distribution of X+H is the summation of two independent random variables, which can be calculated through convolution as follows.



where fX+H(.) and FX+H(.) is the probability density function and cumulative distribution function of X+H respectively. Then the calculation of the expected renewal cycle length is very similar to how we evaluate the expected renewal cycle length for the age-based policy.



For the expected renewal cycle cost, there are three different random events at the end of a renewal cycle to analyze, i.e., a failure, an inspection with replacement due to the defects, and an inspection without replacement. The probability of a failure renewal is the probability that the failure time X+H is before the next inspection comes, i.e.,



Since the failure renewal will incur a corrective maintenance cost Cc, the expected corrective maintenance cost is CcPr(X+H<r). The probability of an inspection renewal without replacement is the probability that the time to defect X is longer than the next inspection time, i.e.,



Since this inspection renewal will only incur an inspection cost Ci, the expected cost related to the inspection renewal without replacement is CiPr(X>r). The probability of an inspection renewal with replacement is the probability that the time to defect X is shorter than the next inspection time and the failure time X+H is longer than the next inspection time, i.e.,



Since this inspection renewal will incur an inspection cost plus the preventive maintenance cost Ci+Cp, the expected cost related to the inspection renewal with replacement is Ci+CpPr(X<r ∩ X+H>r). Therefore, the summation of these cost elements is the expected renewal cycle cost, i.e.,



Notice that the summation of the probabilities of the three renewal events is one. After calculating the expected renewal cycle length and the expected renewal cycle cost, we can obtain the average long run cost rate by ECC/ECL.

To give an example, let us assume both the time to defect and the delay time distributions are exponential with parameters λX 0.6 and λH 0.75 respectively (f(t)= λexp (-λ t)). The time unit is 100 days and the cost parameter values are Cc=1000, Cp=150, and Ci=15 respectively. Using the above equations for ECC and ECL, the expected long run cost per unit time as a function of r is shown in Figure 5. The distribution of X+H is the summation of two independent exponential random variables, which can be calculated through convolution as follows.





example1.png

The optimal inspection interval is 0.4 \* 100=40 days.

A generalized periodic inspection policy for a single-component system

If the time to defect X does not follow an exponential distribution, not all the inspection time points are the renewal points as we discussed in the above section. In this case a renewal cycle may span several inspection intervals. Following the same analysis procedure of renewal theory, we distinguish two renewal events, i.e., the failure renewal and the inspection renewal with replacement. Again, we are interested in the probabilities of the two renewal events. The probability that the failure renewal occurs in a certain inspection interval ((i-1)r,ir) is the probability that the failure happens in the inspection interval ((i-1)r,ir) and the defect occurs after the time point (i-1)r (otherwise the component will be replaced due to the detection of defects), i.e.,



The probability that the inspection renewal with replacement occurs at time point ir is the probability that the failure does not happen till ir and the defect occurs after the time point (i-1)r (otherwise the component will be replaced at (i-1)r), i.e.,



The failure renewals and the inspection renewals can happen at any inspection interval. Notice that the summation of the probability of the failure renewal and the probability of the inspection renewal over all the possible inspection intervals is equal to one,



Then based on the analysis of these two renewal events, the renewal cycle length is



For the renewal cycle cost, when the failure renewal happens in an inspection interval ((i-1)r,ir), the renewal cycle cost equals the sum of the inspection cost (i-1)Ci and the corrective maintenance cost Cc. Then the expected cycle cost due to failure renewals is



When the inspection renewal happens at time point ir, the renewal cycle cost equals the sum of the inspection cost iCi and the preventive maintenance cost Cp. Then the expected cycle cost due to inspection renewals is



Therefore, the expected renewal cycle cost can be obtained by taking the sum of the above two equations,



According to the renewal theory, the expected long run cost rate can be obtained by ECC/ECL.

To give an example, let us assume both the time to defect and the delay time distributions are Weibull with parameters (βX=2, ηX=0.6) and (βH =2, ηH =0.75) respectively. The pdf of Weibull is given as



The time unit is 100 days and the cost parameter values are Cc=1000, Cp=150, and Ci=15 respectively. The expected long run cost per unit time as a function of r is shown in Figure 6. Notice that for Weibull distributions, ECC and ECL don't have close forms. Numerical methods can be used to evaluate the integrals.

example2.png

The optimal inspection interval is 0.3 \* 100=30 days.

Degradation model

If the degradation of a component or a system can be described by a stochastic process over time, we can propose maintenance models based on the probability models of the degradation processes. There are many degradation processes that can be used to describe the changes of the condition of a component or a system. To specify which degradation process is the most appropriate one, we first have to collect information of the failure or degradation mechanism of the component or the system. Hopefully we can get the basic characteristics of the degradation mechanism to screen out the not-appropriate degradation models. The degradation data can be obtained from the reliability testing in labs or from the operating fields. Statistical techniques should be used to estimate the parameters of the degradation models and do the fitting test. Then a suitable degradation model can be selected. In this section, we consider the cases for which the condition of a component can be represented by one variable. Then a stochastic process {X(t),t≥ 0} can describe the degradation process of the condition of a component. If the degradation level X(.) passes a certain failure limit H, as shown in Figure 7, we assume that a failure will happen. This failure can be either a soft failure or a hard failure. We define a soft failure as a failure that will not stop the operation of a system immediately, but will incur extra costs, such as quality loss cost or low performance cost. A hard failure is a failure that will stop the operation of a system immediately.

Degradation1.png

Here are several examples of degrading systems. Let us consider a Micro-Electro-Mechanical System (MEMS) containing one microengine that is subject to wear, see Figure 8. Furthermore, the failure of the microengine causes the failure of the system. The failure of the microengine occurs when the wear volume of material reaches a critical threshold. The wear volume of material can be estimated by measuring the volume of the wear debris or the missing volume in the worn device. For example, a Focused Ion Beam system is effective to evaluate the amount of wear debris by producing cross sections of the precise area of interest in MEMS structures (Peng, Feng, & Coit, 2009).

Degradation2.png

For light display devices, such as plasma display panels (PDPs), vacuum fluorescent displays (VFDs) and fluorescent lamps (FLs), the critical performance characteristic is the luminosity that is related to brightness. Failure of such devices has been traditionally defined in terms of the degradation in luminosity over time. For example, the industry standard definition for PDP lifetime is the time at which the PDP luminosity falls below 50% of its initial luminosity. Failure of FLs is defined as the time when a lamp’s luminosity falls below 60% of its luminosity (Feng, Peng, & Coit, 2010).

A control-limit policy for a single-component system

We consider a CBM policy under which a condition measurement is recorded periodically, and once the measurement is higher than a control limit, repair or replacement of the component may be initiated. Our goal is to optimize the control limit and monitoring interval by minimizing the expected total cost rate. The condition of the monitored item deteriorates over time. Such deterioration may be estimated based on the condition monitoring data that are collected at the monitoring intervals. There are two levels on the degradation process that influence the maintenance actions. The first level is the control limit, denoted by C. Once the deterioration is equal to or higher than this level, the item will be replaced by a new one. The second level is the failure limit H. If the deterioration of the item has reached the failure limit, it will incur a hard failure. This failure limit is usually known and fixed in practice. Upon such failures, the item will also be replaced by a new one. Figure 8 demonstrates the policy. The first passage time of the degradation process {X(t),t≥ 0} over the control limit C is denoted by TC, which can be derived when the degradation process is specified to be a certain stochastic process. The first passage time of the degradation process over the failure limit H is denoted by TH, which can also be derived similarly.

Degradation3.png

We assume the replacement time is negligible compared with the lifetime of the item. If the item gets replaced by a new one, the monitoring schedule restarts after the replacement. Then the replacement time points are renewal points in the analysis. A renewal cycle may span several monitoring intervals. Following the previous analysis procedure of renewal theory, we distinguish two renewal events, i.e., the failure renewal and the inspection renewal when the degradation level is above the control limit but below the failure limit. We are interested in the probabilities of the two renewal events. The probability that the failure renewal occurs in a certain monitoring interval ((i-1)r,ir) is the probability that the degradation level first passes the failure limit in the monitoring interval ((i-1)r,ir) and the degradation level first passes the control limit after the time point (i-1)r (otherwise the component will be replaced according to our CBM policy), i.e.,



where fTC(.) is the probability density function of TC and FTH|TC=u(.) is the cumulative distribution function of TH given that TC=u. The probability that the inspection renewal with replacement occurs at time point ir is the probability that the degradation level does not cross the failure limit till ir and the degradation level crosses the control limit after the time point (i-1)r (otherwise the component will be replaced at (i-1)r), i.e.,



The failure renewals and the inspection renewals can happen at any monitoring interval. Notice that the summation of the probability of the failure renewal and the probability of the inspection renewal over all the possible inspection intervals is equal to one,



Then based on the analysis of these two renewal events, the renewal cycle length is



where fTH|TC=u(.) is the probability density funciton of TH given that TC=u. For the renewal cycle cost, when the failure renewal happens in an inspection interval ((i-1)r,ir), the renewal cycle cost equals the sum of the inspection cost (i-1)Ci and the corrective maintenance cost Cc. Then the expected cycle cost due to failure renewals is



When the inspection renewal happens at time point ir, the renewal cycle cost equals the sum of the inspection cost iCi and the preventive maintenance cost Cp. Then the expected cycle cost due to inspection renewals is



Therefore, the expected renewal cycle cost can be obtained by summing the above two equations,



According to the renewal theory, the expected long run cost rate can be obtained by ECC/ECL.

First passage time: random coefficient model

The above CBM policy requires the derivation of the distributions of first passage time and the condition first passage time. Random coefficient model is one type of probability models to describe the degradation process of a component. In this subsection we will derive the first passage time and the condition first passage time based on a given random coefficient model. For a random coefficient model, the degradation path can be described by a function over time that has fixed coefficients and random coefficients, X(t;Φ,Θ), given a set of constant parameters Φ={φ1, …, φQ}, Q ∈ N and a set of random parameters, Θ={θ1, …, θV}, V ∈ N, following certain probability distributions. The random parameters describe the unit-to-unit variations (instead of the temporal variations). Then clearly, the probability that the first passage time Tχ over a threshold χ is less than time t is equal to the probability that the maximum degradation level in time interval [0,t] is larger than the threshold χ



The conditional distribution of first passage time Tχ2 given that the first passage time of χ1(<χ2) is at time point s is equal to



To give an example, consider a component in a system with a degradation path X(t; Φ,Θ)=φ1+θ1t where Φ={φ1} and Θ={θ1}. The distribution of Tχ can be written in terms of Fθ1 (the cumulative density function of random variable θ1, (θ1 > 0),



Since the randomness is from the unit-to-unit variation, when we know that Tχ1 is s, it is equivalent to say that this component has a degradation rate θ1=(χ1-φ1)/s. Therefore, given that Tχ1 is s, the conditional probability mass function of Tχ2 is



In other words, Tχ2 is a constant number (χ2-φ1)s/(χ1-φ1) given that Tχ1 is s. An illustration of this linear random coefficient model is given in Figure 9.

example3.png

If φ1=1 and θ1 follows Exponential distribution (λ=1), then, when χ1=2 , the first passage time over χ1 is



as shown in Figure 10.

example32.png

If we know the first passage time Tχ1=1 and χ2=3, the conditional probability mass function of Tχ2 given that Tχ1=1 is



In other words, Tχ2 is a constant number 2 given that Tχ1 is 1.

To give an example of how to use it in the inspection optimization model, suppose that the degradation path of a component can be described by a linear random coefficient model. This means X(t)=1+θ1t, θ1 follows Normal(μ=1,σ=0.1) (the negative part is negligible). Suppose the failure limit of this degradation process H is 5. We consider a CBM policy under which a condition measurement is recorded periodically, and once the measurement is higher than a control limit, the component will be replaced by a new one. The inspection cost is Ci=15. The corrective maintenance cost is Cc=1000. The preventive maintenance cost is Cp=150.

example41.png

Thus the expected cycle length based on this linear random coefficient model is



where if (H-φ1)u/(C-φ1)<ir, Pr(TH<ir|TC=u) =1, since the conditional TH is a constant number (H-φ1)u/(C-φ1). By calculating ECC/ECL, the expected long run cost rate as a function of r and C is shown in Figure 11 and 12 (notice that it should be a 3D figure since there are two decision variables but in order to get some insights we choose the 2D figures).

example42.png

*Figure/Table 1. Caption should be sentence case with no ending punctuation if only one sentence*

Please include callouts within the text of your chapter referring to figures and tables to indicate where they should be placed.

All images must be included as separate .tif files, NOT included within the text of your chapter.

Tables should not include cell shading. Any formatting will be removed and will only be presented in black and white.

FuTURE rESEARCH dIRECTIONS

Discuss future and emerging trends. Provide insight about the future of the book’s theme from the perspective of the chapter focus. Viability of a paradigm, model, implementation issues of proposed programs, etc., may be included in this section. If appropriate, suggest future research opportunities within the domain of the topic.

The elementary models for CBM proposed in this chapter are mainly about single-component systems. However, due to the dependencies of components in a system, the models for single-component systems cannot be applied directly on multi-component systems. There are three types of dependencies for multi-component systems, i.e., the economic dependency, the stochastic dependency, and the structural dependency. Relatively less literature have focused on this topic since the models considering these dependencies become more complicated if we want to obtain nice analytical results. It requires more research efforts to deal with this issue in the future.

The development of internet technologies also brings opportunities to further optimize the maintenance scheduling by sharing maintenance resources, e.g., for similar engineering systems located at different physical sites, they can share the maintenance staffs or repair tools. Including this aspect will make the problem more challenging to solve. However, if we are able to come up with an integrated model for maintenance optimization considering the benefits of internet, the operations of maintenance will be performed more efficiently with less cost and higher reliability.

Conclusion

In this chapter, a brief overview of different maintenance strategies is first provided for the readers who are not familiar with maintenance optimization models. The basics of reliability theory are also provided to remind the readers of the fundamental concepts and models. Then several elementary models about CBM are introduced to help the readers get a general idea of the optimization models in this field.

Condition-based maintenance (CBM) is one type of preventive maintenance policy. CBM has attracted lots of attentions of both academia and industry due to the development of advanced sensor technology and measurement devices. The proper implementation of CBM can reduce the frequency of random failures and the expected cost of maintenance during the lifecycle of a system. To obtain the benefits of CBM for companies in capital goods industry, the analytical tools of maintenance optimization are important in the decision processes.

**REFERENCES**

AberdeenGroup. (2003). *Service parts management, unlocking value and profits in the service chain.* Boston: AberdeenGroup.

Christer, A., & Waller, W. (1984). Delay time models of industrial inspection maintenance problems . *Journal of the Operational Research Society* , 401-406.

Ebeling, C. (2009). *An Introduction to Reliability and Maintainability Engineering .* Long Grove: Waveland Pr Inc.

Elwany, A., Gebraeel, N., & Maillart, L. (2011). Structured replacement policies for components with complex degradation processes and dedicated sensors . *Operations Research , 59*, 684–695 .

Feng, Q., Peng, H., & Coit, D. (2010). Joint Optimization of Burn-in, Quality Inspection and Maintenance Policies for Light Display Devices. *International Journal of Advanced Manufacturing Technology* *, 50*, 801-808.

Gebraeel, N., Lawley, M., Li, R., & Ryan, J. (2005). Residual-life distributions from component degradation signals: A Bayesian approach . *IIE Transactions , 37*, 543–557 .

Jasper, J. (2006). *Quick response solution: Fedex critical inventory logistics revitalized.* FedEx: FedEx white paper.

Jardine, A., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance . *Mechanical Systems and Signal Processing , 20*, . 1483–1510 .

Jiang, R., Kim, M., & Makis, V. (2013). Availability maximization under partial observations. *OR Spectrum* *, 35*, 691–710 .

Kharoufeh, J., Solo, C., & Ulukus, M. (2010). Semi-markov models for degradation-based reliability . *IIE Transactions* *, 42*, 599–612 .

Lu, C., & Meeker, W. (1993). Using degradation measures to estimate a time-to-failure distribution . *Technometrics , 35*, 161-174.

Oner, K., Kiesmuller, G., & Van Houtum, G. (2010). Optimization of component reliability in the design phase of capital goods. *European Journal of Operational Research* *, 205*, 615-624.

Peng, H., Feng, Q., & Coit, D. (2009). Simultaneous quality and reliability optimization for microengines subject to degradation . *IEEE Transactions on Reliability* *, 58*, 98-105.

Ross, S. (1996). *Stochastic Processes.* Hoboken: Wiley.

Si, X., Wang, W., Hu, C., & Zhou, D. (2011). Remaining useful life estimation - A review on the statistical data driven approaches. *European Journal of Operational Research , 213*, 1-14.

Tinga, T. (2010). Application of physical failure models to enable usage and load based maintenance. *Reliability Engineering & System Safety* *, 95*, 1061-1075.

Van Noortwijk, J. (2009). A survey of the application of gamma processes in maintenance . *Reliability Engineering & System Safety* *, 94*, 2-21.

Vlok, P., Coetzee, J., Banjevic, D., Jardine, A., & Makis, V. (2002). Optimal component replacement decisions using vibration monitoring and the proportional hazards model . *Journal of The Operational Research Society , 53*, 193–202 .

Wang, W. (2000). A model to determine the optimal critical level and the monitoring intervals in condition based maintenance. *International Journal of Production Research* *, 38*, 1425–1436 .

Key Terms AND Definitions

Please provide 7-10 key terms related to the topic of your chapter and clear, concise definitions (in your own words) for each term. Term and colon should be bold and title case. Definitions should follow a standard dictionary-style format. Place your terms and definitions after the references section of your chapter.

**Appalachia:** A geographic and cultural region of the Mideastern United States. The population in media is portrayed as suspicious, backward, and isolated.

**Ethnocentric:** A belief that one's own culture is superior to other cultures.

**Family-Centricity:** The belief that family is central to well being and that family members and family issues take precedence over other aspects of life.

\*\*\* NOTE: YOUR CHAPTER MUST BE **PROFESSIONALLY COPYEDITED PRIOR TO FINAL SUBMISSION** OF YOUR MANUSCRIPT. FAILURE TO DO SO CAN BE CAUSE FOR EXCLUSION IN THE UPCOMING PUBLICATION. \*\*\*